

DOI:10.3876/j.issn.1000-1980.2013.02.008

基于熵值法的改进集对分析水质模糊评价

郑志宏¹,魏明华²

(1. 华北水利水电学院环境与市政工程学院,河南 郑州 450011; 2. 华北水利水电学院水利学院,河南 郑州 450011)

摘要:为解决水质评价中诸多评价方法在指标权重确定方面存在的困难,应用熵值法改进集对分析模糊评价方法,实现河流水质状况等级评价。该方法将熵值法客观赋权与改进集对分析模糊评价相结合,克服了单纯主观赋权的人为干扰,使评价结果更真实。以滦河下游唐山地区 7 个测站为计算实例,得到各个监测站点的水质评价结果,与综合指数法评价结果符合度为 100%,这些结果与灰色关联法评价结果的符合度为 85.7%,证明了该方法的合理性。

关键词:水质评价;熵值法;改进集对分析;滦河下游;唐山地区地表水

中图分类号:X824 **文献标志码:**A **文章编号:**1000-1980(2013)02-0136-04

Water quality fuzzy evaluation based on entropy method-improved set pair analysis

ZHENG Zhihong¹, WEI Minghua²

(1. Institute of Environmental and Municipal Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450011, China;

2. Institute of Water Conservancy, North China University of Water Resources and Electric Power, Zhengzhou 450011, China)

Abstract: In order to solve the problems in the index weight calculation with multiple methods for water quality evaluation, we used the improved set pair analysis (SPA) fuzzy evaluation method based on information entropy and obtained river water quality estimation grade results. This method combines objective weight and improved SPA fuzzy evaluation and avoids the interference from simple subjective weight, so that the evaluation result can be much more objective. The method was applied to seven monitoring sites in the lower reaches of the Luanhe River in the Tangshan District. The evaluation results of the method showed a conformity degree of 100%, compared with that of the synthetic index method, and a conformity degree of 85.7%, compared with that of the grey relational analysis method, indicating that the proposed method is rational.

Key words: water quality evaluation; entropy method; improved set pair analysis; lower reaches of Luanhe River; surface water of Tangshan District

近年来随着经济的高速发展,不但在水量方面出现了供需矛盾,而且由于各种工业废料、农业化学物质的排放,造成了河流、湖泊、水库等地表水水质严重恶化,直接影响着经济社会的可持续发展。因此,客观评价水质状况关系到人类在内的整个生态系统的健康和安全。目前,水质综合评价方法很多,如层次分析法^[1]、灰色关联法^[2-3]、模糊评判法^[4]、人工神经网络法^[5]、理想区间法^[6]、物元分析法、属性识别法^[7]和主成分分析法^[8]等,各种方法均有其各自的优缺点和使用条件,尤其在评价指标的权重确定方面都存在很大的困难。笔者通过应用基于熵值法的改进集对分析模糊评价方法,以滦河下游唐山地区 7 个代表性测站为研究实例,对滦河下游河流水质状况进行评价,得到符合实际的综合评价结果。

收稿日期: 2012-03-22

基金项目: 水利部公益性行业专项(200801015);华北水利水电学院高层次人才科研启动项目(40214);河南省高校青年骨干教师资助计划(2010GGJS-126);河南省教育厅自然科学研究计划(2011A570009)

作者简介: 郑志宏(1976—),男,河北保定人,副教授,博士,主要从事水资源、水环境系统工程研究。E-mail: zzh@ncwu.edu.cn

1 熵值法原理

在信息论中,信息熵 $H(x) = - \sum p(x_i) \ln p(x_i)$ 反映系统的无序化程度,信息熵越小,系统的无序化程度越大,信息熵越大,系统的无序化程度越小。熵值法就是利用信息熵这一理论工具,根据各项指标值之间的差异程度,计算出各指标的权重,为多指标综合评判问题提供依据。熵值法是一种比较客观的赋权法,这种方法避免了人为因素带来的偏差,并且在众多客观赋权法中是计算较简单的一种^[9]。熵值法的计算步骤如下:

设有 m 个采样点、 n 项指标^[10],构成原始数据矩阵 $\mathbf{X} = (x_{ij})_{m \times n}$ 。

a. 对 x_{ij} 归一化处理:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \tag{1}$$

b. 计算第 j 项指标的熵值:

$$e_j = - \sum_{i=1}^m p_{ij} \ln p_{ij} \tag{2}$$

c. 计算第 j 项指标的权重。对于给定的 j ,指标的权重越大,所揭示的测评指标的差异越大,测评指标的熵值越小,因此令

$$u_j = \frac{1}{e_j} \tag{3}$$

对 u_j 进行归一化处理,使 $\sum w_j = 1$ 。则第 j 项指标的权重 $w_j = u_j / \sum_{j=1}^n u_j$,这样就能得到多目标评判问题各项指标的权重,从而可以利用各种综合评判的方法进行评价。

2 集对分析方法改进

集对分析(set pair analysis, SPA)是赵克勤在 1989 年提出的一种新的系统分析理论方法^[11-12]。该评价方法存在的不足之处为 b, c 的细化问题,同异反评语细化问题。如地表水水质评价中存在 5 类水质标准;一个方案指标的评价语有“好、一般和差”,在好中也有“稍好、好、很好、极好”等细化差别;同样差也有“稍差、差、很差、极差”等细化差别。虽然有学者对这类问题用贴进度来处理^[13],但结果仍不甚理想。笔者对其进行如下改进。

设有 n 个评价指标,将地表水环境指标的各个等级标准 $A_p = \{v_{p1}, \dots, v_{pk}, \dots, v_{pn}\}$ 分别组成一个集合, p 表示等级,将每个区域的地表水环境指标数值系列 $B_q = \{x_{q1}, \dots, x_{qk}, \dots, x_{qn}\}$ 分别组成另一个集合, q 表示区域,这 2 个集合构成一个集对 (A_p, B_q) 。比较 2 个集合中的对应项 x_{qk} 和 v_{pk} ,若 x_{qk} 处于 v_{pk} 的评价级别中,则认为是相同。若 x_{qk} 处于 v_{pk} 的相邻级别中,且在评价级别优越一边,则认为是优异,其值简记为 b_1 ;若 x_{qk} 处于 v_{pk} 的相邻级别中,且在评价级别劣差一边,则认为是劣异,其值简记为 b_2 。若 x_{qk} 处于 v_{pk} 的相隔级别中,且在评价级别优越一边,则认为是优反,其值简记为 c_1 ;若 x_{qk} 处于 v_{pk} 的相隔级别中,且在评价级别劣差一边,则认为是劣反,其值简记为 c_2 。这样, b, c 得到进一步细化,并且这样处理符合原创联系度可展性。具体评价模型见文献^[14]。

3 基于改进集对分析模糊综合评价

模糊综合评价是以模糊数学为基础,对受多种因素制约的不确定性问题进行定量化描述的一种方法。基于集对分析的模糊综合评判的一般步骤如下^[15]:

- a. 确定评价对象的因素(指标)集合 $U = \{u_1, u_2, \dots, u_n\}$,共有 n 个因素。
- b. 确定备择集(或评语集) $V = \{v_1, v_2, \dots, v_m\}$,共有 m 个因素。
- c. 确定权重向量 $\mathbf{W} = [w_1, w_2, \dots, w_n]$ 为 U 中各因素对待评事物的隶属关系。权重集表示各影响因素在评判过程中所占的重要性程度。
- d. 进行单因素评判,建立因素论域和备择论域之间的集对分析联系度矩阵 $\mathbf{R}$ 。

- e. 选择合成算子,将 W 与 R 合成得到 B ,其基本模型为 $B=W \circ R$,“ $\circ$ ”为模糊乘运算算子。
- f. 取集对势最大者为待评方案对应的评判等级。

4 基于熵值法改进集对分析水质模糊评价计算

4.1 数据选取

选取滦河下游唐山地区 7 个监测站点 2003 年的地表水监测指标,分别为冷口(S_1)、崖口(S_2)、邱庄水库(S_3)、小定府庄(S_4)、陡河水库(S_5)、石佛口(S_6)、滦县(S_7)。取溶解氧(x_1)、氨氮(x_2)、高锰酸盐(x_3)、氟化物(x_4)4 项指标作为研究对象,分级标准如表 1。

4.2 熵值法确定指标权重

- a. 建立原始数据矩阵:

$$X = (x_{ij})_{7 \times 4} = \begin{bmatrix} 1.45 & 9.13 & 0.23 & 0.30 \\ 2.62 & 8.27 & 0.33 & 0.60 \\ 5.73 & 8.93 & 0.26 & 0.48 \\ 34.12 & 0.70 & 16.98 & 0.41 \\ 3.66 & 8.63 & 0.13 & 0.63 \\ 2.60 & 10.00 & 0.96 & 0.58 \\ 38.03 & 3.83 & 2.72 & 0.41 \end{bmatrix} \quad (4)$$

表 1 地表水分级标准

Table 1 Surface water quality criterion grades mg/L

等级	x_1	x_2	x_3	x_3
I	≥ 7.5	≤ 0.15	≤ 2	≤ 1.0
II	≥ 6.0	≤ 0.50	≤ 4	≤ 1.0
III	≥ 5.0	≤ 1.00	≤ 6	≤ 1.0
IV	≥ 3.0	≤ 1.50	≤ 10	≤ 1.5
V	≥ 2.0	≤ 2.00	≤ 15	≤ 1.5

- b. 按照式(1) 进行数据转换得到 $X'=(p_{ij})_{7 \times 4}$ 。
- c. 计算第 j 项指标信息熵。
- d. 计算第 j 项指标权重,得到指标权重向量为 $W=[w_1,w_2,w_3,w_4]=[0.178,0.410,0.244,0.168]$ 。

4.3 改进集对分析模糊综合评价

- a. 建立 S_1 监测站点的联系度矩阵:

$$R_1 = \begin{bmatrix} 1 & 0.784 + 0.216i^+ & 0.440 + 0.440i^+ + 0.120j^+ & 0.468 + 0.234i^+ + 0.298j^+ & 0.369 + 0.295i^+ + 0.336j^+ \\ 1 & 0.479 + 0.521i^+ & 0.242 + 0.363i^+ + 0.395j^+ & 0.326 + 0.163i^+ + 0.511j^+ & 0.140 + 0.282i^+ + 0.578j^+ \\ 0.652 + 0.348i^- & 1 & 0.649 + 0.351i^+ & 0.394 + 0.394i^+ + 0.212j^+ & 0.283 + 0.283i^+ + 0.434j^+ \\ 1 & 1i^+ & 1j^+ & 0.417 + 0.583j^+ & 0.417i^+ + 0.583j^+ \end{bmatrix} \quad (5)$$

- b. 应用 Matlab 计算软件对 S_1 点进行模糊运算:

$$B_1 = W \circ R_1 = [0.915 \ 1 + 0.0849i^- \quad 0.580 \ 1 + 0.4199i^+ \quad 0.336 + 0.3127i^+ + 0.3513j^+ \quad 0.383 \ 1 + 0.2046i^+ + 0.4123j^+ \quad 0.192 \ 1 + 0.307i^+ + 0.5009j^+]$$

因此 $shi(A)=a/c$ 或 $shi(A)=a/b$, $shi(u_1)=10.779$, $shi(u_2)=1.382$, $shi(u_3)=0.956$, $shi(u_4)=0.929$, $shi(u_5)=0.384$ 。

分别计算 5 级集对势,得到 S_1 监测站点从 I 级到 V 级的概率分布为:74.7% ,9.6% ,6.6% ,6.4% ,2.7%。所以,得到 S_1 监测站点的评价结果为 I 级。同理,可得到其他 6 个监测站点的评价结果(表 2)。

分别取各监测站点集对势计算结果最大值所在等级为评价结果等级,可得到评价结果如下: S_1,S_2,S_3,S_5 和 S_6 监测站点为 I 级; S_4,S_7 监测站点为 V 级。

5 结 语

使用相同的数据进行计算,综合指数法的评价结果如下: S_1,S_2,S_3,S_5,S_6 监测站点为 I 级; S_4,S_7 监测站点为 V 级。该结果与本文方法评价结果一致,符合度达到 100%。灰色关联法评价结果如下: S_1,S_2,S_3,S_5,S_6 监测站点为 I 级; S_4 监测站点为 V

表 2 各监测站点评价结果

Table 2 Evaluation results of monitoring sites

站点	$shi(u_1)$	$shi(u_2)$	$shi(u_3)$	$shi(u_4)$	$shi(u_5)$
S_1	10.779	1.382	0.956	0.929	0.384
S_2	4.708	2.256	1.414	1.405	0.483
S_3	14.515	1.219	1.410	1.448	0.528
S_4	0.289	0.034	0.033	1.670	9.153
S_5	11.392	1.798	1.346	1.311	0.461
S_6	6.442	0.777	1.155	1.980	0.614
S_7	0.790	0.290	0.468	6.383	7.789

级; S_7 监测站点为Ⅳ级,该结果与本文方法的评价结果符合度达到 85.7%,评价结果只在 S_7 监测站点有略微差异,与其余 6 个监测站点的结果基本一致。以上结果证明了基于熵值法的改进集对分析模糊综合评价模型具有合理性。另外,该方法利用信息熵来确定指标权重,相对来讲更客观,克服了单纯主观赋权方法的人为因素干扰,使评价结果更贴近实际情况,且整个模型计算过程简单便捷,有很好的应用价值。

参考文献:

- [1] 刘章军,叶燎原. 模糊概率模型及其在地下水水质评价中的应用[J]. 应用基础与工程科学学报,2007,15(3):286-293. (LIU Zhangjun, YE Liaoyuan. Fuzzy probability model and its application to evaluation of groundwater quality[J]. Journal of Basic Science and Engineering, 2007,15(3):286-293. (in Chinese))
- [2] 刘思峰,党耀国,方志耕. 灰色系统理论及其应用[M]. 北京:科学出版社,2004:121-126.
- [3] 程诚,李春杰. 灰色关联分析在区域电力市场效率评价中的应用[J]. 华北电力大学学报:社会科学版,2010(3):10-13. (CHENG Cheng, LI Chunjie. Application of grey incidence analysis in the regional electricity market efficiency evaluation[J]. Journal of North China Electric Power University:Social Sciences, 2010(3):10-13. (in Chinese))
- [4] 郑志宏. 郑州市地下水环境评价及修复研究[D]. 郑州:华北水利水电学院,2007.
- [5] 韩立群. 人工神经网络理论、设计及应用[M]. 北京:化学工业出版社,2001:55-78.
- [6] 杨晓华,杨志峰,沈珍瑶,等. 城市环境质量综合评价的多目标决策理想区间法[J]. 系统工程理论与实践,2005,25(8):119-123. (YANG Xiaohua, YANG Zhifeng, SHEN Zhenyao, et al. A multi-objective decision-making ideal interval method for integrated assessment of urban environmental quality[J]. Systems Engineering-theory & Practice, 2005,25(8):119-123. (in Chinese))
- [7] 王硕,张礼兵,金菊良. 系统预测与综合评价方法[M]. 合肥:合肥工业大学出版社,2006:75-98.
- [8] 王静龙. 多元统计分析[M]. 北京:科学出版社,2008:102-136.
- [9] 柯蓉. 熵值法在综合评判股市多指标中的应用研究[J]. 经济论坛,2006(14):107-124. (KE Rong. The application research of entropy method in comprehensive evaluation stock market index[J]. Economic Tribune, 2006(14):107-124. (in Chinese))
- [10] 王元华,曾凤章. 基于熵值法的顾客满意度测评[J]. 商业研究,2004(22):11-13. (WANG Yuanhua, ZENG Fengzhang. An evaluation method of customer satisfaction degree based on information entropy[J]. Commercial Research,2004(22):11-13. (in Chinese))
- [11] 赵克勤. 集对分析及其初步应用[M]. 杭州:浙江科学技术出版社,2000:1-3.
- [12] 邓红霞,李存军,朱兵,等. 基于集对分析的生态承载能力综合评价方法[J]. 长江科学院院报,2006,23(6):35-38. (DENG Hongxia, LI Cunjun, ZHU Bing, et al. Integrative assessment of eco-carrying capacity based on set pair analysis[J]. Journal of Yangtze River Scientific Research Institute, 2006,23(6):35-38. (in Chinese))
- [13] 周超明,晏路明. 港湾水质综合评价的集对分析模型[J]. 环境科学与管理,2006,31(8):62-66. (ZHOU Chaoming, YAN Luming. A module of set pair analysis on the comprehensive assessment of seawater quality of harbors and bays in Fujian Coast [J]. Environmental Science and Management, 2006,31(8):62-66. (in Chinese))
- [14] 魏明华,郑志宏,黄强,等. 基于改进 SPA 法的地下水环境模糊综合评判[J]. 水利学报,2009,40(10):1204-1209. (WEI Minghua, ZHENG Zhihong, HUANG Qiang, et al. Fuzzy comprehensive evaluation of groundwater environment based on improved set pair analysis[J]. Journal of Hydraulic Engineering, 2009,40(10):1204-1209. (in Chinese))
- [15] 姜军,宋保维,潘光,等. 基于集对分析的模糊综合评判[J]. 西北工业大学学报,2007,25(3):421-424. (JIANG Jun, SONG Baowei, PAN Guang, et al. A novel method for fuzzy comprehensive evaluation based on set pair analysis[J]. Journal of Northwestern Polytechnical University, 2007,25(3):421-424. (in Chinese))